

REMARKS

Reconsideration and withdrawal of all grounds of rejection are respectfully requested in view of the above amendments and the following remarks. Claims 1-15 are now pending.

The Examiner had previously objected to the title of the application as being non-descriptive and it was amended in response to the first office action. The title is amended again to add reference to the fact that apparatus claims are pending in the application and that the application concerns both apparatus and method.

One aspect of the present invention is an ability to quickly identify a random subset of the data within a large database. A data mining procedure can then be run against the data subset instead of the whole data set to produce meaningful results at a fraction of the time needed to run the analysis on the entire database. The invention seeks to solve a need in the art as discussed on page 1, line 26 to page 2, line 2 of the application.

Although many algorithms for such problems are known and widely used (for example, Decision Trees and K-Means Clustering), they take too much time if trained on too much data. It has been observed that under certain circumstances, however, it may not be necessary to use an entire database (which can have many millions of records) to create a useful model or predictor. Instead a sample of a few tens of thousand records might accurately represent the much larger data set of the entire database.

Independent claim 1 of the present invention is directed to a computer-implemented process. Claim 1 features a method of identifying a subset of records within a database for purposes of representing the database. The method is accomplished by choosing a selection attribute from one of a plurality of attributes contained in records within the database and scanning records in the database. A randomizing function is applied to the selection attribute of records that are scanned to create a randomized record value. A selection criteria identifies records for inclusion within a subset of records. This is done by comparing the randomized record value of each record to which the randomizing function was applied with the selection criteria.

Claim 1 now stands rejected under 35 USC 103 based on the combined teaching of US patents no. 6,510,427 to Bossemeyer, Jr. et al (herein Bossemeyer) in view of US patent no. 5,960,431 to Choy. The patent to Bossemeyer concerns searching over an entire database to

select a subset of the records of the database. Applicant's invention provides a representation of the database with a random selection of records from the database by applying a randomizing function to the records of the database over a so-called selection attribute and then comparing the result of that randomizing function with a selection criteria. Bossemeyer teaches away from this feature of the invention by suggesting use of an attribute (without randomizing function being applied) to select a subset of records. This is identical to any selection search over an entire database for selection of records whose attributes satisfy a condition.

As an example, consider a typical SQL query such as "SELECT * from TABLE A where A.weight >= 200 " This will cause a database management system to find all records from Table A where the weight attribute is greater than or equal to 200. This is consistent with the statement at column 5 of Bossemeyer "The search engine 60 responds by searching the database 16 for all records meeting the search criteria ... the records 42 meeting the subject search criteria are compiled and displayed in a search result 70" See Bossemeyer col 5, lines 43 – 54. One would use the Bossemeyer system with the above representative selection SQL query to retrieve all records from the table A of individuals having a weight greater than 200 pounds. This result, however, would be a poor representation of the entire database. Bossemeyer does not recognize nor satisfy a goal of applicants invention, i.e providing a data subset that represents the database. The result of that search would have not individuals having a weight less than 200 pounds and hence would be a poor representation. For this reason, it seems apparent that Bossemeyer teaches away from applicants invention featured in claim 1. There would be no motivation to combine the system of Bossemeyer with the teaching of Choy since the goal of Bossemeyer is so fundamentally different from the subject matter of claim 1.

Turning now to the Choy patent, one sees a reference in the Background art section of this patent to a transformation of the value of one or more data fields to divide a database up into sections for record access. There is no suggestion in this patent, however, for the process of applying a selection criteria to reduce the size of the database and produce a representative sample of the database. For this reason, even if the combination of references proposed by the examiner is valid, the combination does not teach all elements of the claim and therefore does not provide a prima facie basis of rejection. For at least these reasons claim 1 is patentable.

Claims 2 – 6 depend from allowable claim 1 and are allowable.

Independent claim 7 of the application is directed to a client/server computer apparatus.

One or more client computers are coupled to a network and include communications instructions for requesting a data set from a server computer by means of the network. The server computer is coupled to the network and has access to a database having a number of records. The server computer includes instructions for sending a dataset made up of a subset of the records in the database to a client computer via the network. The server computer includes instructions for scanning records in the database, applying a randomizing function to a specified record attribute of each scanned record in the database to produce a randomized record value, and comparing the randomized record value with a selection criteria to determine whether to include a record in the subset of records from the database for transmission via the network to the client.

All arguments presented above with regard to claim 1 apply to apparatus claim 7. There is no motivation to combine the Bossemeyer and Choy patents in the manner proposed by the Examiner since the Bossemeyer teaches away from selection process that provides a smaller data set than an original but still provides a representation of the original. This is buttressed by the fact that the combination is lacking an element, i.e. the use of a selection criteria to determine whether a record is to be included in the data subset and therefore the rejection does not provide a prima facie grounds of rejection.

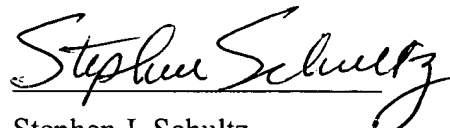
Claims 8 and 9 depend from allowable claim 7 and are also allowable.

Independent claim 10 is directed to a computer readable medium including instructions for identifying a subset of records within a database for purposes of representing the database. In short, the claim recites instructions for performing the steps recited in claim 1. Therefore, for at least the same reasons as provided above in regard to the patentability of claim 1, claim 10 is allowable. Further, it is submitted that claims 11-15 are allowable at least by virtue of dependence on allowable claim 10. Consequently, withdrawal of this rejection is respectfully requested.

In view of the above, it is respectfully submitted that the invention of independent claims 1, 7 and 10 is patentable. Further, the subject matter of the remaining dependent claims is patentable at least by virtue of direct or indirect dependence on claims 1, 7 and 10. Therefore, it is believed that all pending claims of this application are in condition for allowance. Accordingly, entry of the Amendment and a subsequent early Notice of Allowance for all pending claims of this application is respectfully solicited.

Respectfully submitted,

Date: March 1, 2004
Telephone: (216) 241-6700
Facsimile: (216) 241-8151


Stephen J. Schultz
Reg. No. 29,108